대한수학교육학회지 <학교수학> 제 9 권 제 2 호
Journal of Korea Society of Educational Studies in Mathematics
School Mathematics Vol.9, No.2, 223-240. Jun. 2007

평가 문항을 활용한 중학교 수학 교육과정의 내용 및 인지행동의 위계성 조사

송 미 영* · 김 선 희**

본 연구는 중학생들의 수학 성취를 국가수준에서 평가한 경험적 자료를 활용하여 우리나라 중학교 수학과 교육과정의 내용과 수학에서의 인지행동이 위계적으로 구성되어 있는지를 조사하였다. 전반적으로 교육과정의 내용 제시 순서는 난이도 순위와 통계적으로 유의한 상관관계가 나타나지 않은 반면, 인지행동의 위계는 난이도 순위와 통계적으로 유의한 상관관계가 있었다. 이러한 결과에서 검사 문항의 난이도 순위가 학교에서 배운 수학 교과 내용의 순서보다는 문항에서 요구하는 인지행동의 수준과 더 관련이 있음을 알 수 있었다. 그리고 내용 위계와 인지행동의 위계 간 상관관계가 유의하게 나타나, 교육과정에서 늦게 등장하는 내용일수록 요구되는 인지행동도 높은 수준임을 발견할 수 있었다. 내용 및 인지행동의 위계와 난이도 순위 간 상관분석에서 특이한 양상을 나타낸 문항에 대해서는 그 특성을 분석하였다.

Ⅰ. 서 론

교육과정의 질에 대한 평가는 학생들이 반드시 성취해야 할 내용이 포함되었는지, 그 내용이 학생들의 인지 수준에 적정한지, 내용 간의 위계가 적절하고 학습하기에 용이한 순서로 조직되었는지, 실현 가능한지, 전체적으로 일관성 있게 구성되었는지, 정확하고 원활하게 의사소통 되는지 등 다면적으로 이루어질 수 있다(박선화 외, 2006). 특히 학생들에게 쉬운 내용이 어려운 내용에 앞서 도입되어야 하는 것이 적절하다는 측면에서(NCTM, 1989), 학생들이 학습하기에 용이한 순서로 수학과의 교육내용이 배열되어 있는지, 즉 수학 교육과정의 내용이 수학의 논리적 체계 속에서 학년에 따라 위계적으로 구성되어 있는지에 대한 검토는 필요하다.

학교 교육의 결과로서 학생들이 얻게 되는 교육적 경험은 학교 내에서 교사와 학생의 상호 작용 속에서 실행된 교육과정에 의존하게 되고, 학교 내에서 교육과정의 실행은 교육과정 문서나 교과서 등에 드러난 국가 교육과정을 기초로 하여 이루어진다. 따라서 학생들의 수학 성취를 평가한 경험적 자료를 통해 수학과 교육과정의 특성에 대한 실증적 검토는 교육과정의 질 판단을 위한 의미 있는 기초 자료를 산출할 수 있다. 한 예로 Shermis & Chang (1997)은 미국 내 한 대학의 수학 교과목의 이수 내용 순서와 진단평가 문항의 난이도 순위 사이에 유의한 상관관계가 있음을 밝혀, 대학 수학 교육과정에 대하여 위계성이 적절함을 보인 바 있다.

우리나라 수학과 교육과정의 질에 대한 평가는 교육과정 개정 시에 집중적으로 이루어지

* 한국교육과정평가원 부연구위원(mysong@kice.re.kr)
** 한국교육과정평가원 부연구위원(math1207@kice.re.kr)

며, 학교 현장의 수학 교사들과 수학, 수학교육 전문가들의 의견을 중심으로 현대 수학교육의 동향이 반영되어 새로운 대안을 모색하는 양상을 띠고 있다. 즉 학습의 주체인 학생들의 입장에서 교육과정에 대한 평가는 이루어지지 않고 있으며, 보다 쉬운 것이 어려운 것에 선행되어 구성되었는지에 대한 실증적인 검토는 거의 이루어지지 않고 있다. 이에 본 연구는 교육과정의 범위를 중학교 교육과정에 한정하여 국가수준에서 실시된 학업성취도 평가 자료를 통해 우리나라 수학과 교육과정의 내용 구성이 위계적인 본질을 갖고 있는지 파악해 보고자 한다. 교육과정의 질을 평가하는 일은 다양한 측면에서 이루어져야 할 것이나 본 연구에서는 학생의 성취라는 실증적 자료를 통해 문항의 난이도로 교육과정 내용의 위계를 탐색하려는 것이다. 이 때 학생의 성취는 교과 지식 내용뿐 아니라 인지행동에서도 일어나므로, 두 영역에서의 위계성을 점검할 것이다.

수학과 교육과정 문서의 진술에는 교과 지식인 내용영역뿐 아니라 인지행동도 포함되어 있다. 예를 들면 '…를 이해하고 이를 구할 수 있다.'와 같이 내용과 인지행동이 조합된 형식으로 학생들이 성취해야 할 기준이 진술되어 있다. 이렇게 진술된 교육과정을 기초로 학생이 얻게 되는 교육적 경험은 교과 지식 내용뿐 아니라 인지행동도 포함된다. 같은 수준의 동일 내용영역은 여러 인지행동과 연관지어 학습될 수 있으며, 더 고차원적인 능력 및 기술을 익힐 때 쉬운 주제로 다루어질 수도 있다.

따라서 본 연구는 중학교 수학과 교육과정에서 학생들이 성취하기를 기대하는 교과 지식 내용과 수학에서의 인지행동, 그리고 그 위계에 대한 고찰을 토대로 교육과정 문서에서 제시된 교과 내용의 조직 위계 및 인지행동의 위계와 국가수준 학업성취도 평가 문항의 난이도 순위 간 상관관계를 파악하고자 한다. 이를 통하여 중학교 수학과 교육과정의 내용과 인지행동의 위계성을 점검하고, 어떤 것이 문항의 난이도와 더 강한 상관관계를 갖는지, 내용 조직의 순서와 인지행동의 수준은 상관이 있는지 분석하고자 한다. 이러한 목적을 달성하기 위하여 본 연구에서 설정한 연구 문제는 다음과 같다.

첫째, 중학교 수학 교육과정의 내용영역 위계와 문항난이도 순위 간에 상관관계가 있는가?

둘째, 수학 교과의 인지행동 위계와 문항난이도 순위 간에 상관관계가 있는가?

셋째, 중학교 수학 교육과정의 내용영역 위계와 인지행동 위계 간에 상관관계가 있는가?

II. 인지행동의 위계

수학 교육을 통해 학생들은 다양한 수학적 능력을 신장시킬 기회를 갖는다. 기본적인 알고리즘에 따라 수나 문자를 다루는 계산 능력, 수학적 개념을 알고 그것을 적용할 수 있는 이해 능력, 논리나 발견을 조장할 수 있는 추론 능력, 문제 상황에서 수학적 지식을 전략적으로 활용할 수 있는 문제해결 능력 등은 수학 교육을 통해 가장 잘 형성될 수 있다. 이러한 수학적 능력은 인지행동으로도 불리며, 여러 평가 틀에서 행동 영역으로도 분류되고 있다.

예를 들어, 경제협력개발기구(Organization for Economic Co-operation and Development; OECD)의 학업성취도 국제 비교 연구인 PISA (Programme for International Student Assessment)는 수학적 능력을 수학적 사고와 추론, 수학적 논쟁, 수학적 의사소통, 모델링, 문제 제기와 문제해결, 표현, 상징적 · 형식적 · 기법적인 언

어와 조작의 활용, 보조 교구와 도구의 활용 등의 8가지로 제시한다. 이러한 능력들은 많은 부분에서 중첩되며 '실제 상황'에서 수학을 할 때는 다양한 능력들이 동시에 필요하기 때문에, 학교 교육을 받은 학생들이 장차 사회에 나가서 생산적인 역할을 할 준비가 되어 있는가를 점검하기 위한 연구인 PISA는 각각의 능력을 포괄하여 재생군(reproduction cluster), 연결군(connection cluster), 반성군(reflection cluster)으로 구분하고 있다(이미경 외, 2004). PISA 2000 (노국향, 최승현, 박경미, 2001)에서 이 세 가지의 능력군을 수준 1, 2, 3으로 표현한 것을 볼 때, 각각의 능력군에 수준의 위계가 내포되어 있음을 알 수 있다.

국제교육평가협회(International Association for the Evaluation of Educational Achievement: IEA)가 추진하는 수학·과학 추이 변화 국제비교 연구인 TIMSS(Trends in Mathematics and Science Study) 평가 틀에도 인지행동이 제시되어 있다. 학생들이 수학 내용을 학습함에 따라 습득하기를 기대하는 인지행동은, TIMSS 2007에서는 알기(knowing), 적용하기(applying), 추론하기(reasoning)의 세 영역으로 구성된다(Mullis et al, 2005). 이는 TIMSS 2003에서 사실과 절차 알기, 개념의 활용, 정형적인 문제해결, 추론의 네 영역으로 구분된 인지행동이 수정된 것이다.

우리나라 국가수준 학업성취도 평가[1]에서는 수학과 관련된 인지행동을 황혜정·최승현(1999)의 연구에 따라 계산, 이해, 문제해결, 추론, 의사소통으로 나눈다. '계산'은 여러 가지 계산 방법뿐 아니라 문제해결에 이르기 위한 명확한 절차, 즉 알고리즘을 능숙하게 구사할 수 있는 능력에 관한 것이고, '이해'는 기본적인 수학적 개념, 원리, 법칙 및 그 관련성을 이해하여 의미 충실한 개념적 사고를 형성할 수 있는 능력에 관한 것이다. '추론'은 관찰, 열거, 실험 등을 통한 귀납, 유추, 추측에 의해 수학적 법칙과 문제의 해법을 발견할 수 있는 능력, 조건명제의 증명, 삼단논법에 의한 연역적 추론, 반례에 의한 증명, 간접증명법, 모순법, 동치인 명제의 증명, 수학적 귀납법 등을 이용한 증명을 읽고 이해할 수 있으며 이러한 방법을 사용하여 수학적 명제를 증명할 수 있는 능력을 말한다. '문제해결'은 수학의 여러 가지 내용 사이의 개념, 원리, 법칙 등의 관련성이 요구되는 수학 내적인 문제를 해결할 수 있는 능력과, 수학과 일상생활 및 타교과와 관련성이 있는 통합교과적인 소재의 응용 문제를 해결할 수 있는 능력에 관한 것이다. '의사소통'은 계산, 개념, 추론, 문제해결 영역에 관한 문제를 해결하는 상황에서 주어진 문제 상황과 관련된 수학적 내용을 토대로 수학적 용어, 기호, 문장 등을 이용하여 그 해결 과정의 근거와 이유를 표현할 수 있는 능력을 말한다. 의사소통 능력은 구성형, 개방형의 문항에서 적절하게 평가될 수 있으며 선다형에서는 측정하기 어렵다.

위의 여러 평가 틀에 따른 인지행동의 구분을 볼 때, 인지행동은 평가의 목적과 문항 유형에 따라 다름을 알 수 있다. 그리고 하나의 평가 틀 내에서의 인지행동 구분도 규정하기에 따라 수정 가능하다. 또한 하나의 문항이 어떤 능력 한 가지만을 측정한다고 규정하기에 어려움이 있다는 것도 PISA의 능력군 개념을 통해 알 수 있다.

인지행동의 위계를 설정하기 위해 본 연구는 성취도 평가 문항 각각에 인지행동을 대응시키

---

1) 이하 성취도 평가라 함.

고 그 위계를 규정하였다. 인지행동의 위계를 규정하는 것은 크게 두 가지 차원에서 이루어질 수 있다. 첫째는 평가 틀의 인지행동 간의 위계이고, 둘째는 각 인지행동 내에서의 위계이다.

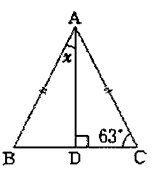
첫 번째, 인지행동 간의 위계는 PISA의 능력군 개념을 적용하여 그 위계를 생각해 볼 수 있다. 본 연구는 성취도 평가의 선다형 문항을 사용할 것이므로, 그 평가 틀에 따라 인지행동을 계산, 이해, 추론, 문제해결로 구분한다. '계산'은 기본적인 알고리즘을 구사할 수 있는 재생군 수준의 능력으로, 다른 인지행동에 비해 낮은 수준이라 할 수 있다. '이해'는 단순한 지식의 기억보다는 그 관련성을 이해하는 것으로 연결군에 해당한다고 할 수 있다. 그리고 '추론'과 '문제해결'은 논증이나 일반화, 문제의 수학화 등에 해당하는 반성군이다. 이때 추론은 문제해결에서도 적용되는 능력이므로, 문제해결을 추론보다 더 상위 수준으로 두었다. 따라서 인지행동은 '계산-이해-추론-문제해결'의 순으로 높은 수준이 되는 위계라 할 수 있다.

두 번째, 각 인지행동 내에서의 위계는 복잡성의 정도에 따라 그 위계를 정할 수 있다. 동일한 '계산'이라도 주어진 식의 복잡성에 따라 수준이 다를 수 있으며, '이해' 내에서도 관련 개념의 수준이나 얼마나 많은 개념의 연결이 필요한지에 따라 그 수준이 다를 수 있다. 또한 이것은 '추론'이나 '문제해결'에서도 마찬가지이다. Nesher, Hershkovitz & Novotna(2003)는 세 양의 비교를 S-P(공유된 부분) 스키마와 H(위계적) 스키마, S-W(공유된 전체)의 스키마로 나누고 이 스키마 각각에서의 복잡성 지수를 도출하여, 문제에서 요구되는 과정이 제시되는 조건에 따라 얼마나 복잡한지를 구하였다. 스키마별로 문제해결에 사용되는 복잡성 지수는 그 문제를 해결하는 데 있어 학생들이 얼마나 많은 개념과 절차를 활용해야 하는지를 나타낸다. 이 방법을 적용한다면, 동일한 인지행동 내에서 복잡한 정도에 따라 문항의 위계성을 세워, 각 문항간 상대적인 복잡성으로 인지행동 내에서의 순서를 정할 수 있다.

이러한 방식에서 의하여 인지행동의 위계를 설정하는 구체적인 예는 다음과 같다. [그림 Ⅱ-1]과 [그림 Ⅱ-2]는 모두 도형 영역에서 문제해결 행동을 요구하는 문항이다. [그림 Ⅱ-1]에 제시된 문항을 풀기 위해서 학생들은 △ABD와 △ACD이 합동임을 파악하고 삼각형의 내각의 합이 180°임을 이용하여 x의 크기를 구하여야 한다. 이에 비해 [그림 Ⅱ-2]는 닮음의 위치에 있는 두 사각형에서 닮음의 중심이 무엇인지 파악하고, "닮음의 중심으로부터 두 도형 위의 대응하는 점까지의 거리의 비가 일정하고 그것이 닮음비와 같다"는 성질을 생각하여, 주어진 변의 길이로 어떻게 닮음비를 만들어야 하는지 찾아야 한다. [그림 Ⅱ-1]에서 합동인 삼각형을 시각적으로 바로 인지하고 각의 크기도 쉽게 구할 수 있는 반면, [그림 Ⅱ-2]에서는 관련된 성질을 적용하기에 도형이 더 복잡하고 비례식을 이용하는 계산이 더 복잡하므로 [그림 Ⅱ-1]보다 높은 수준의 문제해결 행동이 요구된다고 할 수 있다. 즉, 동일한 인지행동 내에서도 문항의 복잡성에 따라 그 위계를 설정하는 것이다.

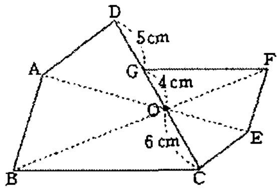
4. $\overline{AB}=\overline{AC}$인 이등변삼각형 ABC에서 점 D는 점 A에서 변 BC에 내린 수선의 발이다. ∠ACD = 63° 일 때, ∠BAD의 크기는?

① 23° ② 27° ③ 30°
④ 33° ⑤ 37°



[그림 Ⅱ-1] 2005년 성취도 평가 A형 4번 문항

12. 사각형 ABCD와 사각형 EFGC가 서로 닮음의 위치에 있다. $\overline{GO} = 4\text{cm}$, $\overline{OC} = 6\text{cm}$, $\overline{DG} = 5\text{cm}$일 때, 두 도형의 닮음비는?



① 3 : 2 ② 4 : 3 ③ 5 : 4
④ 6 : 5 ⑤ 7 : 6

[그림 II-2] 2005년 성취도 평가 A형 12번 문항

III. 연구 방법

1. 연구 자료

본 연구에서는 중학교 수학과 7차 교육과정의 내용과 인지행동의 위계성을 분석하기 위해서 2004년과 2005년 성취도 평가의 중학교 3학년용 수학 검사와 이에 대한 응답 자료를 사용하였다. 성취도 평가는 교육과정에서 규정하고 있는 교육목표에 비추어 학생의 목표 도달도, 교육과정의 문제점과 정착 정도, 연도별 변화 추이 등을 파악하기 위하여 매년 10월 시행되는 전국 규모의 평가이다. 성취도 평가의 중학교 3학년용 수학 검사는 중학교 교육과정의 전체 6개 내용영역, 즉 수와 연산, 도형, 측정, 확률과 통계, 문자와 식, 규칙성과 함수에서 계산, 이해, 추론, 문제 해결, 의사소통의 인지행동을 측정하는 문항 36개로 구성된다.

본 연구는 현재까지 성취도 평가 결과가 발표된 최근 2개년, 즉 2004년 성취도 평가의 중학교 3학년용 수학 검사와 2005년 성취도 평가의 중학교 3학년용 수학 검사 중 A형[2]의 선다형 문항을 분석 대상으로 선택하였다. 이 검사도구는 선다형 30문항과 구성형(constructed-response item) 6문항으로 구성된 혼합형 검사이다. 혼합형 검사는 각 문항 유형간 특성의 차이로 인한 유형 효과(format effect)가 발생할 수 있다. 선다형 문항과 구성형 문항이 동일한 영역과 내용을 다루도록 제작하더라도 각 문항 유형이 측정하는 특성(trait), 수험자가 문항을 인지하는 과정과 정서적 반응 등에 차이가 있다고 보고되었다(노국향, 김신영, 2000; Birenbaum & Feldman, 1998; Thissen, Wainer, & Wang, 1994; Zeinder, 1987). 이러한 문항 유형의 차이를 고려하여 본 연구는 각 문항 유형에서 요구되는 인지행동이 차별화되기 때문에 선다형 문항만을 선택하여 분석하였다. 따라서 본 연구에서 분석한 문항의 수는 2004년 검사 중 선다형 30문항과 2005년 A형 검사 중 선다형 30문항, 총 60개의 선다형 문항이다. 각 문항의 내용영역과 인지행동은 [부록]에 제시되어 있으며, <표 III-1>은 내용영역과 인지행동별 문항 수를 보여준다.

<표 III-1> 내용 및 행동영역별 문항 분포

내용 \ 행동	계산	이해	추론	문제 해결	합계
수와 연산	2	4	1	2	9
도형	0	3	5	4	12
측정	1	3	1	3	8
확률과 통계	1	3	1	5	10
문자와 식	7	2	0	3	12
규칙성과 함수	0	6	1	2	9
합계	11	21	9	19	60

2) 2005년 성취도 평가 이후에는 변화추이 분석을 위한 검사점수 동등화의 목적으로 교과별 검사 도구 2종(A형, B형)을 개발·시행하고, 이 중 B형 검사의 문항은 공개하지 않음.

2004년과 2005년 성취도 평가는 전국의 중학교 3학년에 재학 중인 학생 중 1%의 학생을 표집하여 당해연도 10월 셋째 주에 실시되었다(조영미, 이대현, 이봉주, 2004; 이양락 외, 2005). 시·도교육청, 도시화, 학교 규모 등을 고려한 2단계 층화군집표집설계(A two-stage stratified cluster sample design)에 의해 평가대상으로 선정되어 2004년 검사를 치른 6,276명의 문항반응자료와 2005년 A형 검사를 치른 3,309명의 문항반응자료를 분석하였다.

2. 분석 방법 및 절차

I 장에서 기술한 연구문제에 따라 성취도 평가 문항의 내용영역 위계 및 인지행동 위계를 각각 설정하고, 이들과 문항난이도간 상관관계를 산출하였으며, 특이한 경향을 보이는 문항 내용에 대한 질적 분석을 수행하는 일련의 절차를 거쳐 연구가 진행되었다.

가. 내용영역의 위계 설정

성취도 평가 중학교 3학년용 검사도구에 포함된 선다형 60문항이 측정하는 내용영역의 위계는 7차 교육과정과 교과서에 제시되어 있는 순서에 따라 설정하였다. 교과서를 편찬할 때 교육과정을 재구성할 수 있도록 되어 있으나 7차 교육과정이 적용된 교과서들은 내용 순서를 거의 동일하게 따르고 있다. 따라서 총 60개 문항의 내용영역 위계를 교육과정과 교과서에서 제시되는 순서에 따라 정하고, 6개 내용영역 내에서도 이와 같은 방식에 의하여 그 위계를 정하였다. <표 Ⅲ-2>는 60개 선다형 문항의 주된 평가 내용을 6개 내용영역 및 단계별로 간략하게 제시한 것이며, 평가 내용은 교육과

<표 Ⅲ-2> 내용영역별 평가 내용

내용 영역	단계	평가 내용
수와 연산	7	집합, 자연수의 성질(소인수분해, 최대공약수, 최소공배수), 십진법과 이진법, 정수와 유리수의 개념과 대소관계, 정수와 유리수의 사칙계산
	8	유리수와 소수, 유리수와 순환소수
	9	제곱근과 실수(무리수의 개념, 실수의 대소 관계와 수직선), 근호를 포함한 식의 계산
도형	7	기본도형, 작도와 합동, 평면도형의 성질, 입체도형의 성질
	8	삼각형과 사각형의 성질, 도형의 닮음, 닮음의 응용,
	9	피타고라스의 정리 및 활용
측정	7	다각형과 각의 크기, 도형의 길이, 넓이, 부피
	8	근사값과 오차, 근사값의 덧셈과 뺄셈
확률과 통계	7	도수분포와 그래프, 상대도수의 분포와 누적도수의 분포
	8	확률과 그 기본성질
	9	상관도와 상관표
문자와 식	7	문자의 사용과 식의 계산, 일차방정식, 일차방정식의 활용
	8	식의 계산, 미지수가 2개인 연립일차방정식, 연립일차방정식의 활용, 일차부등식과 연립일차부등식, 일차부등식과 연립일차부등식의 활용
	9	다항식의 곱셈과 인수분해, 이차방정식, 이차방정식의 활용
규칙성과 함수	7	함수와 그 그래프(정비례, 반비례, 제1, 2, 3, 4사분면)
	8	일차함수와 그 그래프, 일차함수의 활용(평행이동)
	9	이차함수와 그 그래프(축, 꼭지점, 최대값, 최소값)

정에 근거하여 작성된 것이다.

나. 행동영역의 위계 설정

성취도 평가의 평가 틀에 따라 인지행동은 계산, 이해, 추론, 문제해결의 4개 영역으로 구분하였다. 행동영역에 대한 위계 설정은 성취도 평가 문항 검토에 참여한 수학교육 전문가 3인의 의견을 토대로 이루어졌다. 60개의 선다형 문항을 제시하고, 계산, 이해, 추론, 문제해결의 개념과 성격을 연구자가 설명하였고, 4가지 인지행동의 위계를 Ⅱ장에서와 같이 계산-이해-추론-문제해결의 순으로 합의하고 각 문항이 어떤 인지행동에 초점을 둔 것인지 구분하였다. 또한, 각 행동영역 내에서는 제시된 인지행동의 복잡성에 따라 그 위계를 설정하였다. 복잡성의 정도는 Nesher 등(2003)의 연구를 참조하여 문제에 제시된 식을 간단히 하고, 관련 개념이 여러 개이고, 식을 세우고, 답을 도출하고, 해석하는 등에서의 복잡성으로 평가하였다. 각 문제마다의 복잡성은 도출하지 않고 상대적인 복잡성으로 인지행동 내에서의 순서를 정하였다. 이와 같은 방법으로 전문가들의 합의에 따라 각 문항의 행동영역과 그 위계가 설정되었다.

다. 문항난이도의 추정

문항의 쉽고 어려운 정도를 나타내는 난이도 지수는 대개 전체 응답자 중 정답자의 비율로 추정한다. 이는 고전검사이론에 의한 문항난이도 지수로서 검사를 치루는 집단의 능력 수준에 따라 다르게 추정될 수 있다. 하지만 문항반응이론에 의하여 추정된 문항난이도는 집단의 특성에 관계없이 동일하다는 특징이 있다. 성취도 평가에서는 두 검사이론을 모두 적용하고 있으며[3], 문항반응이론에 의해 분석할 때 선다형 문항에는 3모수 로지스틱 모형을 적용하여 문항변별도와 난이도 및 추측도를 추정한다. 3모수 로지스틱 문항반응모형에서 학생 i가 선다형 문항 j에 정답을 맞힐 확률 P_{ij}는 다음과 같이 표현된다.

$$P_{ij} = c_j + (1 - c_j)\frac{\exp[Da_j(\theta_i - b_j)]}{1 + \exp[Da_j(\theta_i - b_j)]}$$

$$(\text{단, } i = 1, 2, \cdots, G,\ j = 1, 2, \cdots, K)$$

이 식에서 a_j는 문항 j의 변별도, b_j는 문항 j의 난이도(위치모수), c_j는 문항 j의 추측도, θ_i는 학생 i의 능력모수를 가리키며, D는 1.7이다. 고전검사이론에 의한 정답률은 0부터 1사이의 값을 가지며 값이 클수록 쉬운 문항으로 해석하지만, 문항반응이론에 의한 위치모수는 이론적으로 음의 무한대부터 양의 무한대 사이의 값을 가지며 값이 클수록 어려운 문항으로 해석된다. 본 연구에서 분석대상으로 선택한 검사는 2004년 성취도 평가와 2005년 성취도 평가의 검사로서, 각 검사를 치른 집단의 수준이 동일하다고 보장할 수 없기 때문에 집단의 능력 수준에 영향을 받는 정답률 이외에 문항반응이론에 의한 위치모수도 사용하였다. 이때 위치모수는 연도간 검사유형간 검사점수 비교를 위한 동등화(equating) 작업으로 공통척도화한 것이다.

성취도 평가에서 활용되는 동등화 기법은 문항반응이론에 의한 진점수 동등화 방법이며, 이 방법의 핵심이 되는 모수의 공통척도화를 간략히 개관하면 다음과 같다(박정 외, 2006). 우선 기존 검사(예: 2004년 검사)의 응답 자료와 새로운 검사(예: 2005년 A형 검사)의 응답 자료에 대해 각각 문항반응모형의 문

3) 성취도 평가에서 두 검사이론을 적용하는 목적이 상이함. 고전검사이론은 문항의 양호도 분석을 위해서 사용되고, 동등화를 위해서는 문항반응이론이 사용됨.

항모수와 능력모수를 추정한다. 문항반응이론의 모수 불변성 특성으로 인해 각 모수치는 선형으로 연결되어 있다(Lord, 1980)는 속성을 이용하여, 새로운 검사의 문항모수와 능력모수 추정치를 기존 검사의 척도, 즉 공통척도로 변환한다.

본 연구에서 사용할 문항난이도의 추정치, 즉 문항반응이론에 의한 위치모수와 고전검사이론에 의한 정답률간 등위상관계수를 산출한 결과 .925(p=.000)로 매우 높은 상관을 갖고 있는 것으로 나타났다. 이러한 결과로부터 공통척도화된 위치모수의 순위와 정답률의 순위는 크게 다르지 않으며, 난이도 순위를 비교하는 경우에는 위치모수와 정답률 중 어느 것을 사용해도 무방하다고 하겠다. 60개 선다형 문항의 정답률은 [부록]에 제시되어 있다.

라. 문항의 위계와 난이도의 관계 파악

교육과정과 교과서에 제시되어 있는 순서에 따라 설정한 문항의 내용영역 위계, 수학 교육 전문가의 합의에 따라 설정한 인지행동 위계가 각각 문항난이도 순위와 어느 정도 관련성이 있는지를 분석하기 위하여 Spearman의 등위상관계수(rank correlation coefficient)를 산출하고 상관도(scatter plot)를 그렸다. 난이도 순위는 쉬운 문항에서 어려운 문항으로 순서를 배열하였다. 내용영역 위계와 인지행동 위계, 문항난이도 순위는 서열 척도에 의한 변수이므로 다음 식에 의해 계산되는 Spearman의 등위상관계수 r_s과 그 유의도 검정에 의해 각 변수간 상관성을 파악하였다.

$$r_s = 1 - \frac{6\sum D_i^2}{n(n^2-1)}$$

이 식에서 D_i는 대상 i의 두 등위 간 차이, n은 사례수를 가리킨다. 등위상관계수를 산출할 때에는 문항난이도 지수로서 문항반응이론의 위치모수와 고전검사이론의 정답률을 모두 사용하였고, 상관도를 그릴 때에는 위치모수보다 보편적으로 널리 사용되고 있는 정답률을 사용하였다.

Spearman의 등위상관계수 r_s의 값은 -1에서부터 +1 사이의 값을 가지며, 그 절대값이 클수록 강한 상관관계, 양의 부호이면 정적인 상관관계, 음의 부호이면 부적인 상관관계인 것으로 해석할 수 있다. 상관계수에 대한 유의도 검정을 위한 영가설은 '모집단의 상관계수는 0이다.', 유의수준은 .05로 설정하였다. 본 연구 자료에서 산출된 상관계수에 대한 유의확률이 유의수준 .05보다 작으면 두 변수 간에 통계적으로 유의한 상관관계가 있다고 해석한다. 예를 들어, 교육과정 내용 위계와 문항난이도 순위 간 등위상관계수가 통계적으로 유의하다면 문항난이도를 지표로 할 때 교육과정 내용은 위계적이라고 볼 수 있는 것이다.

마. 특이 문항 분석

내용영역 위계 및 인지행동 위계와 문항난이도의 상관도에서 선형성을 크게 벗어나는 문항을 추출하여, 어떠한 이유에서 특이성을 보이는지를 파악하기 위해 그 문항에서 요구하는 내용영역과 인지행동을 살펴보았다.

Ⅳ. 연구 결과

1. 내용영역의 위계성

내용영역의 위계가 존재하는지 분석하기 위하여 교육과정 내용 제시 순서와 문항난이도 순위간 Spearman의 등위상관계수를 산출한 결과는 <표 Ⅳ-1>과 같다.

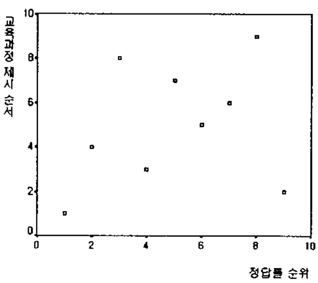
<표 Ⅳ-1> 내용영역 위계와 문항난이도 순위간 등위 상관계수

내용영역	위치모수	정답률
수와 연산 (9문항)	.133 (.732)	.283 (.460)
문자와 식 (12문항)	.315 (.319)	.406 (.191)
규칙성과 함수 (9문항)	.567 (.112)	.650 (.058)
확률과 통계 (10문항)	.503 (.138)	.212 (.558)
도형 (12문항)	-.245 (.443)	-.077 (.812)
측정 (8문항)	-.619 (.102)	-.214 (.610)
전체 (60문항)	.113 (.390)	.147 (.262)

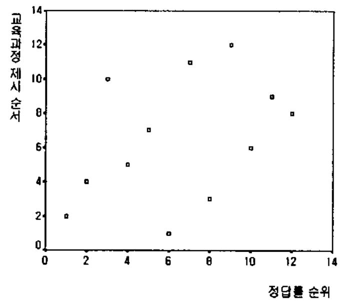
* (　) 안은 유의확률

총 60개 문항의 교육과정 내용 제시 순서와 문항반응이론의 위치모수 순위에 대해 등위상관계수를 구한 결과 통계적으로 유의한 결과는 아니었다(r=.113, p=.390). 교육과정 내용 제시 순서와 고전검사이론에서의 정답률 순위 사이의 상관계수도 통계적으로 유의한 결과는 나타나지 않았다(r=.147, p=.262). 중학교 수학 교육과정 전체로는 내용의 위계성이 문항난이도 순위와 상관관계가 있지 않았다.

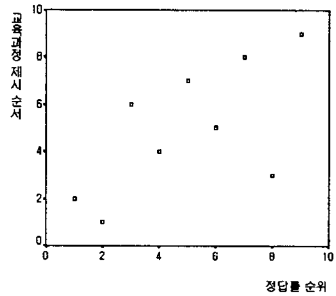
각 내용영역 내에서 위계가 존재하는지 살펴보면, 모든 영역에서 교육과정 내용 제시 순서와 위치모수 순위 간의 상관관계, 정답률 순위 간의 상관관계가 통계적으로 유의하지 않았다. 이러한 경향은 [그림 Ⅳ-1]~[그림 Ⅳ-6]에 제시된 내용 위계와 정답률 순위간 상관도에서도 관찰된다.



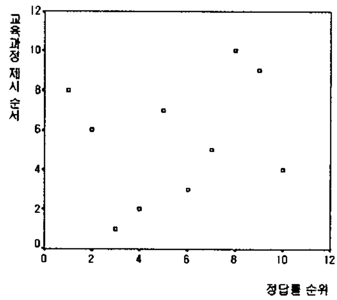
[그림 Ⅳ-1] 수와 연산 영역에서 내용 위계와 정답률 순위와의 상관도



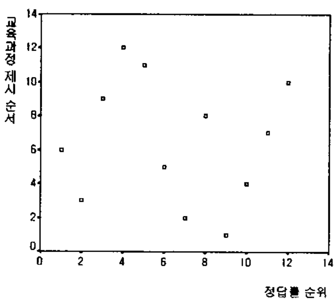
[그림 Ⅳ-2] 문자와 식 영역에서 내용 위계와 정답률 순위와의 상관도



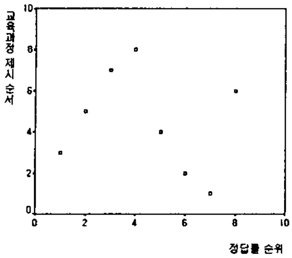
[그림 Ⅳ-3] 규칙성과 함수 영역에서 내용 위계와 정답률 순위와의 상관도



[그림 Ⅳ-4] 확률과 통계 영역에서 내용 위계와 정답률 순위와의 상관도



[그림 Ⅳ-5] 도형 영역에서 내용 위계와 정답률 순위와의 상관도



[그림 Ⅳ-6] 측정 영역에서 내용 위계와 정답률 순위와의 상관도

'규칙성과 함수' 영역의 9개 문항은 내용 제시 순서와 정답률 순위와의 등위상관계수가 .650(p=.058)으로 다른 영역에 비해 가장 높았고, '규칙성과 함수' 영역의 교육과정 제시 순서는 성취도 평가의 문항 정답률 순위와 상관관계가 있다고 볼 수도 있다. '규칙성과 함수' 영역은 중학교 수학 교육과정에서 정비례와 반비례를 통하여 함수의 개념을 도입하고, 정비례와 반비례 그래프를 학습하며, 일차함수, 이차함수의 내용으로 전개되는데, 이 내용 전개의 순서가 문항의 정답률 순위와 상관관계가 있는 것이라 하겠다.

2. 인지행동의 위계성

본 연구에서 설정한 인지행동의 위계와 문항난이도 순위 간 Spearman의 등위상관계수를 산출한 결과는 <표 Ⅳ-2>와 같다.

<표 Ⅳ-2> 행동영역 위계와 문항난이도 순위간 등위상관계수

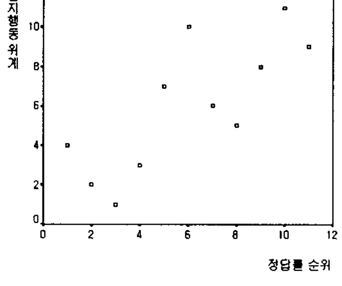
인지행동	위치모수	정답률
계산 (11문항)	.664 (.026)	.773 (.005)
이해 (19문항)	.478 (.028)	.491 (.024)
추론 (21문항)	.683 (.042)	.167 (.668)
문제해결 (9문항)	.526 (.021)	.556 (.013)
전체 (60문항)	.440 (.000)	.406 (.001)

* (　) 안은 유의확률

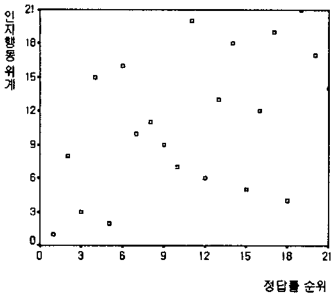
인지행동 위계와 위치모수의 순위 및 정답률 순위 사이의 등위상관계수는 각각 .440, .406이었으며 통계적으로 유의하였다(p=.000; p=.001). 이것은 인지행동의 위계가 문항난이도 순위와 서로 관련 있다는 것을 말해 준다. 특히 인지행동의 위계를 계산, 이해, 추론, 문제해결의 순서로 설정한 것이 적절하였음을 문항난이도 순위와의 상관관계를 통해 알 수 있다.

4가지 인지행동에서의 위계성을 각각 살펴보면, 인지행동 내에서의 위계와 위치모수 순위 간에는 모두 통계적으로 유의한 상관관계가 있었으며, 등위상관계수의 범위는 .478~.664이었다. 인지행동 위계와 정답률 순위 간에는 추론 영역을 제외하고는 모두 통계적으로 유의한 상관관계가 있었으며 등위상관계수의 범위는 .406~.773이었다. 이러한 결과가 나타난 것은 일반화나 논증 등의 추론은 문제의 복잡성보다는 추론 능력 자체의 위계로 그 순서가 정해질 수 있기 때문으로 사료된다.

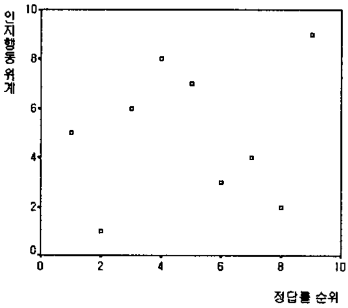
[그림 Ⅳ-7]~[그림 Ⅳ-10]은 정답률 순위와 인지행동 위계 간의 상관도이다. 각 행동영역 내에서 문제의 복잡성이 높을수록 문항의 정답률 또한 낮아짐을 알 수 있다. 이와 같은 분석 결과를 볼 때, 본 연구에서 설정한 인지행동의 위계는 문항난이도의 순위와 상호 관련이 있다. 인지행동의 수준이 높을수록 그 문항은 어려운 경향을 보인다고 할 수 있다.



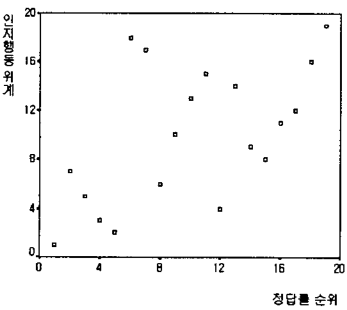
[그림 Ⅳ-7] 계산 영역에서 인지행동 위계와 정답률 순위와의 상관도



[그림 Ⅳ-8] 이해 영역에서 인지행동 위계와 정답률 순위와의 상관도



[그림 Ⅳ-9] 추론 영역에서 인지행동 위계와 정답률 순위와의 상관도



[그림 Ⅳ-10] 문제해결 영역에서 행동 위계와 정답률 순위와의 상관도

3. 내용영역과 인지행동 위계의 상관관계

교육과정의 내용영역 위계와 인지행동 위계가 상호 관련이 있는지 알아보았다. <표 Ⅳ-3>에 제시된 바와 같이 교육과정 내용 위계와 인지행동 위계 사이에는 통계적으로 유의한 상관관계가 있었다(r=.371, p=.003). 전반적으로 교육과정에서 늦게 등장하는 내용일수록 요구되는 인지행동도 높은 수준이라 할 수 있다.

<표 Ⅳ-3> 내용영역 위계와 행동영역 위계 간 등위 상관계수

구분	영역 (문항 수)	상관계수 (유의확률)
내용 영역	수와 연산 (9문항)	.550 (.125)
	문자와 식 (12문항)	.559 (.059)
	규칙성과 함수 (9문항)	.400 (.286)
	확률과 통계 (10문항)	.903 (.000)
	도형 (12문항)	.266 (.404)
	측정 (8문항)	-.595 (.120)
행동 영역	계산 (11문항)	.255 (.450)
	이해 (19문항)	-.025 (.915)
	추론 (21문항)	.583 (.099)
	문제해결 (9문항)	.067 (.786)
전체 (60문항)		.371 (.003)

내용영역별로 상관계수를 살펴보면, '확률과 통계' 영역에서의 내용 제시 순서와 행동영역의 위계와의 상관계수가 .903(p=.000)으로 상당히 높은 편이었다. '확률과 통계' 영역의 내용 제시 순서는 계산, 이해, 추론, 문제해결의 위계에 따라 조직되는 것과 관련이 있다고 할 수 있다. 하지만 문항의 수가 충분히 많지는 않기 때문에 '확률과 통계' 교육과정이 아닌 성취도 평가의 문항이 높은 학년에서 다루는 내용일수록 높은 수준의 인지행동을 사용하도록 구성된 것은 아닌지 살펴볼 필요는 있다.

한편, 인지행동별로 교육과정 내용 위계와의 상관계수를 조사하였을 때는 통계적으로 유의한 결과를 얻지 못했다. 이는 동일한 인지행동이 모든 내용 영역을 통틀어 반복적으로 교수·학습되도록 조직되어 있음을 반증하는 결과라고도 할 수 있다.

4. 특이 문항 분석

이 절에서는 교육과정의 위계와 문항난이도 순위 간에 통계적으로 유의한 상관관계가 있다고 판명된 일부 영역에서 관찰된 특이 문항의 특성을 분석하였다. 규칙성과 함수 영역은 교육과정 내용 위계가 정답률 순위와 상관이 있다고도 할 수 있었는데, <그림 Ⅳ-3>을 보면 몇 개의 문항은 교육과정 내용 위계와 정답률 순위의 선형성에서 벗어나고 있다. 또한 추론을 제외하고는 인지행동 위계와 정답률 순위 간에 유의한 상관관계가 있었는데, <그림 Ⅳ-8>과 <그림 Ⅳ-10>에서 보듯이 이해와 문제해결 영역에서 선형성에서 벗어난 몇 개의 문항들이 있다. 선형적인 관계에서 가장 멀리 떨어져 있는 문항들을 특이 문항으로 선정하고 그 문항들을 분석해 보고자 한다.

가. 규칙성과 함수 영역

규칙성과 함수 영역에서 선정된 특이 문항은 2004년 검사의 21번 문항으로, 이 문항은 일차함수의 기울기에 대한 것이며, 인지행동으로는 이해에 해당한다. 규칙성과 함수 영역에서 출제된 9개의 문항 중에서 교육과정 내용 순서상 3번째에 해당하는 것으로 초기에 도입되는 내용이지만, 정답률 순위로는 8번째로 낮은 편이다. 기울기의 의미만 안다면 풀 수 있는 그리 복잡하지 않은 문항으로 판단되며 인지행동 '이해'에 해당하는 19문항 중에서도 순위가 4위였다. 개념에 대한 이해가 단편적으로 기억될 때 인지행동이 낮은 수준의 복잡하지 않은 것이라 할지라도 정답률은 낮을 수 있음을 보여주는 것이며, 중학교 3학년 학생들에게 일차함수의 기울기 개념이 그러하다고 볼 수 있다.

21. 기울기가 $\frac{3}{2}$인 일차함수의 그래프에서 x의 값이 1에서 5까지 증가할 때, 다음 중 y의 값의 변화를 바르게 설명한 것은?

① 3만큼 증가한다.
② 4만큼 증가한다.
③ 4만큼 감소한다.
④ 6만큼 증가한다.
⑤ 6만큼 감소한다.

내용영역	규칙성과 함수		인지행동	이해
답지 반응분포(%)	①	②	③	④	⑤
	9.9	24.4	17.5	41.3	6.8

* 답지반응분포에서 음영처리된 것이 정답임.

나. 이해 영역

이해 영역에서의 특이 문항인 2005년 검사의 17번 문항은, 19개의 이해 영역 문항 중 5순위로 복잡성이 상대적으로 낮다. 하지만 정답률이 51.5%로 이해 영역 내에서 15번째로 정답률이 낮았다. 즉, 복잡성이 낮고 인지행동 중 낮은 수준인 이해 문제이지만 학생들은 정답을 잘 맞히지 못한 것이다. 답지반응분포에서 28.2%의 학생들이 답지 ①을 선택한 것을 볼 때, 포물선의 폭이 이차항의 계수와 관련이 있다는 것은 알지만, 계수의 절대값이 작을수록 폭이 넓어진다는 것을 반대로 이해하였음을 알 수 있다. 즉, 일부 학생들은 포물선의 폭이 가장 넓을 때와 가장 좁을 때의 값을 혼동하고 있었다. 복잡한 이해 능력을 요구하지 않았지만, 학생들은 포물선의 폭에 대한 잘못된 내용을 기억하고 있어서 문항의 정답률이 낮은 것으로 보인다. 일차함수의 기울기에 대한 2004년 검사의 21번 문항의 반응과 마찬가지로 도구적 이해의 수준에서 포물선의 폭에 대한 개념이 형성되어 나타난 결과라 하겠다.

17. 다음 포물선 중 폭이 가장 넓은 것은?

① $y=-3x^2$

② $y=-2x^2$

③ $y=2x^2$

④ $y=x^2$

⑤ $y=\frac{1}{3}x^2$

내용영역	규칙성과 함수		인지행동	이해
답지 반응분포(%)	①	②	③	④	⑤
	28.2	3.7	12.1	4.3	51.5

다. 문제해결 영역

문제해결 영역에서의 특이 문항으로는 2005년에 출제된 16번 문항을 살펴보았다. 문제해결 영역에 속한 19개의 문항 중에서 순위가 18인 복잡한 문제였으나 정답률 순위는 6위로서 학생들에게 상대적으로 쉬운 문제였다고 할 수 있다. 이 문제를 해결하기 위해서는 직사각형의 가로나 세로의 길이를 미지수로 놓고 식을 구해야 한다. 가로의 길이를 x라 한다면 세로의 길이가 $x-2$이고, 변형된 텃밭의 넓이는 $(x+3)(x-2-1)=16$가 되어 이차방정식을 풀어야 한다. $x=\pm5$가 되고 문항에서 x는 길이이므로 $x=5$라 하면, 처음 텃밭의 넓이는 $5\times3=15\text{m}^2$이 된다. 복잡한 문제해결을 해야 했지만 학생들의 정답률은 상당히 높은 것으로 나타났다. 직사각형의 넓이를 이용하는 이차방정식의 활용 문제는 상당히 정형화된 것으로 교과서에서 많이 등장하고 있다. 중학교 3학년 학생에게 친숙하고 학습한 지 오래 되지 않은 내용이고, 문제를 해결하는 과정에서 만든 이차방정식이 간단하여 정답률이 높았을 수도 있다. 실생활 문제에 학생들이 취약하다는 것을 생각해 볼 때, 실생활 문제라고 해서 학생들이 모두 어려워하는 것은 아님을 알 수 있다. 이에 대한 더 심층적인 조사가 앞으로 이루어져야 할 것이다.

16. 가로가 세로보다 2m 긴 직사각형 모양의 텃밭이 있다. 이 텃밭의 가로를 3m 늘리고 세로를 1m 줄였더니 넓이가 16m^2가 되었다. 처음 텃밭의 넓이는?

① 15m^2 ② 16m^2 ③ 20m^2

④ 24m^2 ⑤ 25m^2

내용영역	문자와 식		인지행동	문제해결
답지반응분포 (%)	①	②	③	④	⑤
	63.5	9.3	12.3	10.6	4.0

V. 결론 및 제언

본 연구는 성취도 평가 자료를 활용하여 중학교 수학 교육과정의 내용 제시가 위계적으로 조직되었는지와 수학에서의 인지행동이 위계적인지 알아보고자 성취도 평가 문항의 난이도 순위와의 상관관계를 조사하였다. 교육과정의 내용 제시 순서는 전반적으로 문항 정답률 순위와 통계적으로 유의한 상관관계가 나타나지 않았다. 내용영역의 순서와 문항난이도 순위간에 상관관계가 나타나지 않은 것은 내용영역의 위계보다 인지행동의 위계가 난이도를 결정하는 데 더 중요한 요인일 수 있지만, 중학교 3학년 학생들에게 내용 위계상 높은 순서인 것이 최근에 배운 내용으로 더 잘 알고 있는 내용일 가능성도 있으므로 이에 대해서는 더 심도 있는 분석과 논의가 필요하다.

인지행동은 성취도 평가 틀의 4개 영역, 즉 '계산-이해-추론-문제해결'의 순으로 위계를 설정한 후 문제의 복잡성에 따라 영역 내에서의 위계를 설정하여 문항난이도 순위와의 등위상관계수를 통해 위계성을 알아보았다. 행동영역

에서 전체적인 위계는 문항난이도 순위와 상관관계가 있었다. 이러한 결과는 성취도 평가 문항의 난이도가 문항에서 요구하는 인지행동의 수준과 관련이 있음을 말해준다 하겠다. 각 행동영역 내에서의 위계와 문항난이도 순위와의 상관관계를 조사하였을 때에도 통계적으로 유의한 결과가 나타나, 행동영역 내에서 문제의 복잡성이 문항난이도 순위와 관련이 있음을 알 수 있다. 그러나 행동영역 중 '추론'은 정답률 순위와 상관계수가 통계적으로 유의하지 않았는데, 이것은 '추론' 영역에서는 문항의 복잡성보다는 추론 능력 자체의 수준에 따라 문항난이도가 결정될 수도 있음을 시사한다.

수학과 교육과정 내용의 제시 순서와 인지행동의 위계 간에 상관관계가 있는지 조사한 결과, 전체적으로 통계적으로 유의한 상관관계가 있었고, 영역별로는 '확률과 통계'에서 상관계수가 높았으며 통계적으로 유의하였다. 즉, 학년이 높을수록 보다 높은 수준의 인지행동을 요하는 내용을 다루며, 특히 '확률과 통계' 영역에서 그러한 경향이 강하다고 할 수 있다.

본 연구에서 제시한 기준에 따라 설정된 내용이나 인지행동의 위계와 문항난이도 순위와의 상관관계를 분석하는 과정에서 일반적인 경향인 선형성을 많이 벗어난 특이 문항이 관찰되었으며, 그 이유를 탐색하기 위하여 특이 문항의 특성을 분석하였다. 각 문항을 분석한 결과 교육 내용의 순서나 인지행동의 위계와 별도로 학생들이 개념을 잘못 이해하거나 암기하는 성향을 갖고 있다면 그러한 내용이나 인지행동을 다루는 문항은 일반적인 경향인 선형성에서 이탈됨을 알 수 있었다. 본 연구에서는 일차함수 기울기의 정의에 따른 의미와 포물선의 폭과 관련된 내용이 그러한 것으로 파악되었다. 학생들이 무엇을 모르고 잘못 알고 있는지에 대하여 파악하고자 할 때도 본 연구의 방법을 이용하여 특이한 사항을 파악하는 것은 유용할 것이며, 앞으로 보다 많은 개념적 오류를 파악해 나가야 할 것이다. 또한 특이 문항 분석을 통해 실생활 문제이고 복잡한 문제해결 과정이 도입되더라도 학생들이 모두 어려워하고 기피하는 것은 아니라는 것을 알 수 있었다. 실생활 상황을 수학화하고 문제를 해결할 수 있는 능력을 기르는 것은 수학 교육의 목표 중 하나이며, 학생들이 이러한 문제에 어떤 반응을 보이는지에 대한 분석이 더 심층적으로 이루어져야 할 것이다.

본 연구에서는 중학생들의 수학 성취 결과가 반영된 경험적 자료를 활용하여 중학교 수학 교육과정 구성의 위계성을 점검하는 실증적 사례를 제시하였다는 점에서 의의가 있으나, 중학교 수학 교육과정에 국한하여 위계를 살펴보았다. 앞으로 초등학교나 고등학교 과정에 대한 질 판단의 일환으로 이와 같은 연구를 확장해 볼 수 있을 것이다. 특히 고등학교에서는 학생들이 과목을 선택할 수 있는 체계로 되어 있으므로 수학 Ⅰ이나 Ⅱ를 이수한 후 어떤 선택과목이 위계성이 있는지 확인해 볼 수도 있을 것이다.

우리나라는 국가수준의 교육과정이 구성되고 학교 교육과정에서 실현되기 때문에 본 연구에서는 국가수준 학업 성취도 평가 결과를 활용하였지만, 교육과정의 모든 내용이 빠짐없이 평가되지는 못했다는 한계가 있다. 또한 문항 유형의 효과를 배제하기 위하여 선다형 문항만 선택하여 분석함으로써, 일부 내용에 해당하는 인지행동이 고루 갖추어지지 못했다는 한계가 있다. 후속 연구에서는 모든 내용이 포함될 수 있도록 충분한 수의 문항으로 구성된 검사를 활용하여 대규모로 학생의 성취도를 조사하고 그 결과를 분석해 볼 것을 제안한다. 그리고 본 연구에서는 수학 교육과정의 위계에 대하여 문

항난이도라는 관점에서만 탐색했으나 다측면적으로 평가하려는 시도가 계속 필요할 것이다.

참고문헌

노국향 · 김신영(2000). 문항의 형태에 따른 피험자의 인지적 · 정의적 반응의 차이에 관한 연구. **교육평가연구, 13**(1), 181-194.

노국향 · 최승현 · 박경미(2001). **PISA 2000 수학 평가 결과 분석 연구**. 한국교육과정평가원 연구보고 RRE 2001-9-3.

박선화 · 고정화 · 권점례 · 김선희 · 도종훈 · 신성균 · 최승현 · 나귀수 · 박경미 · 이만근 · 정순영 · 정영옥 · 권혁천 · 전성진 · 최동렬(2006). **수학과 교육과정 개정 시안 수정 · 보완 연구**. 한국교육과정평가원 연구보고 CRC 2006-9.

박정 · 김경희 · 김수진 · 손원숙 · 송미영 · 조지민(2006). **국가수준 학업성취도 평가 -기술보고서-**. 한국교육과정평가원 연구보고 RRO 2006-4.

이미경 · 곽영순 · 민경석 · 채선희 · 최성연 · 최미숙 · 나귀수(2004). **PISA 2003 결과 분석 연구 - 수학적 소양, 읽기 소양, 과학적 소양 수준 및 배경변인 분석-**. 한국교육과정평가원 연구보고 RRE 2004-2-1.

이양락 · 김선희 · 고정화 · 조영미 · 구자형(2005). **2004년 국가수준 학업성취도 평가 연구 -수학-**. 한국교육과정평가원 연구보고 RRE 2005-1-4.

조영미 · 이대현 · 이봉주(2004). **2003년 국가수준 학업성취도 평가 연구 -수학-**. 한국교육과정평가원 연구보고 RRE 2004-1-4.

황혜정 · 최승현(1999). 수학과 평가틀에 관한 고찰. **수학교육학연구 9**(2), 459-471.

Birenbaum, M., & Feldman, R. A.(1998). Relationships between learning patterns and attitudes towards two assessment formats. *Educational Research, 40*, 90-98.

Lord, F. M.(1980). *Applications of Item Response Theory to practical testing problems.* Hillsdales, NJ: Erlbaum.

Mullis, I.V.S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., Arora, A., & Eberber, E.(2005). *TIMSS 2007 Assessment Frameworks.* IEA.

NCTM(1989). *Curriculum and Evaluation standards for school mathematics.* Reston, VA: Author.

Nesher, P., Hershkovitz, S., & Novotna, J. (2003). Situation model, text base, and what else? Factors affecting problem solving. *Educational Studies in Mathematics, 52*, 151-176.

Shermis, M.D. & Chang, S. (1997). The use of item response theory(IRT) to investigate the hierarchical nature of a college mathematics curriculum. *Educational and Psychological Measurement, 57*(3), 450-458.

Thissen, D., Wainer, H. & Wang, X. B. (1994). Are tests comprising both multiple-choice and free-response items necessarily less unidimensional than multiple-choice test: An analysis of two test. *Journal of Educational Measurement, 31*, 11-123.

Zeinder, M.(1987). Essay versus multiple-choice type classroom exams: the student's perspective, *Journal of Educational Research, 80*, 352-358.

Investigating the Hierarchical Nature of Content and Cognitive Domains in the Mathematics Curriculum for Korean Middle School Students via Assessment Items

Song, Mi Young (KICE)
Kim, Sun Hee (KICE)

The purpose of this study was to investigate the degree to which the middle school mathematics curriculum matched the item difficulty levels of representative mathematics items. The items used in this study were developed for the National Assessment of Educational Achievement.

Ranks for difficulty values of the 60 multiple-choice item were calculated via both Classical Test Theory and Item Response Theory and correlated with the rank order of the mathematics content and cognitive domains sequence. There are six content domains; number and operation, algebra, measurement, figure, pattern and function, and probability and statistics. The cognitive domains include computation, understanding, reasoning and problem-solving.

Results suggest a congruence between cognitive domain's sequence and item difficulty levels of items based on that sequence. This finding indicates that the linear or hierarchical assumptions concerning the sequence appears to be reasonable. The characteristics of items that were exceptions to this trend were addressed.

* key words : Hierarchical Nature(위계성), Mathematics Curriculum(수학 교육과정), Cognitive domain(인지행동), Item Difficulty(문항난이도)

논문접수 : 2007. 4. 30
심사완료 : 2007. 5. 25

[부록] 문항 정보

연도	문항 번호	내용영역	행동영역	정답률
2004	1	수와 연산	이해	90.4
2004	2	수와 연산	계산	77.2
2004	3	도형	이해	91.4
2004	4	도형	추론	54.0
2004	5	수와 연산	이해	52.5
2004	6	측정	이해	75.9
2004	7	확률과 통계	문제해결	55.4
2004	8	확률과 통계	문제해결	79.5
2004	9	문자와 식	계산	84.8
2004	10	문자와 식	이해	72.0
2004	11	규칙성과 함수	이해	56.5
2004	12	규칙성과 함수	이해	55.8
2004	13	도형	추론	50.5
2004	14	측정	문제해결	80.7
2004	15	도형	추론	50.1
2004	16	측정	계산	67.8
2004	17	확률과 통계	이해	71.7
2004	18	문자와 식	계산	57.9
2004	19	수와 연산	이해	32.9
2004	20	문자와 식	계산	64.8
2004	21	규칙성과 함수	이해	41.3
2004	22	규칙성과 함수	추론	59.5
2004	23	도형	문제해결	61.4
2004	24	규칙성과 함수	문제해결	47.7
2004	25	확률과 통계	추론	33.5
2004	26	도형	이해	51.4
2004	27	문자와 식	문제해결	26.1
2004	28	수와 연산	문제해결	47.4
2004	29	확률과 통계	문제해결	50.5
2004	30	도형	추론	28.3
2005	1	문자와 식	계산	85.6
2005	2	규칙성과 함수	이해	81.9
2005	3	확률과 통계	문제해결	73.7
2005	4	도형	문제해결	92.1
2005	5	수와 연산	이해	76.8
2005	6	문자와 식	이해	69.0
2005	7	문자와 식	계산	71.7
2005	8	문자와 식	계산	64.0
2005	9	수와 연산	계산	57.1
2005	10	수와 연산	문제해결	49.0
2005	11	확률과 통계	이해	67.6
2005	12	도형	문제해결	81.7
2005	13	측정	추론	73.1
2005	14	문자와 식	계산	70.0
2005	15	확률과 통계	문제해결	52.6

2005	16	문자와 식	문제해결	63.5
2005	17	규칙성과 함수	이해	51.5
2005	18	규칙성과 함수	이해	46.2
2005	19	도형	이해	64.2
2005	20	수와 연산	추론	52.5
2005	21	확률과 통계	이해	53.8
2005	22	도형	추론	54.9
2005	23	도형	문제해결	54.2
2005	24	측정	이해	39.3
2005	25	측정	문제해결	59.0
2005	26	측정	문제해결	41.3
2005	27	측정	이해	16.1
2005	28	문자와 식	문제해결	42.6
2005	29	확률과 통계	계산	55.0
2005	30	규칙성과 함수	문제해결	36.9